Analysis using Belief Propagation Techniques

A Report on Course Work Seminar
Submitted in partial fulfillment of the requirement of

University of Mumbai

For the Degree of

Ph.D
(Electronic Engineering)

by

Chitra Suresh

Under the guidance of

Prof. Dr.Kushal.R.Tuckley

RAIT

Department of Electronics Engineering
Ramrao Adik Institute of Technology,
Sector 7, Nerul , Navi Mumbai
(Affiliated to University of Mumbai)
December 2014

Abstract

Probabilistic models have received considerable attention,particularly as they are cater for uncertain expert knowledge in their representation and reasoning. The knowledge in these models is represented by causal relationships between variables and their probability measures.The alternate Bayesian approach uses the degree of a person's belief that an event will occur,rather than the actual probability that the event will occur.

The graphical representation of the dependencies embedded in the probabilistic models.The graphical representation used for belief propagation Techniques are two types.They are Markov Random Fields and Bayesian Networks. Belief Propagation Techniques are used for performing inference on graphical models such as Bayesian Network and Markov Random Field.

Belief Propagation Techniques are used in wide variety of algorithms which are used in artificial intelligence,signal processing and digital communication.

Keywords:Belief Propagation Technique,Bayesian Network and Markov Random Field

Contents

List of Figures

Chapter 1

Introduction

Probabilistic models have received considerable attention,particularly as they are cater for uncertain expert knowledge in their representation and reasoning. The knowledge in these models is represented by causal relationships between variables and their probability measures.These probability measures may be based on classical [16] probabilities,Bayesian probabilities or combination of both.
Classical probabilities relies upon repeated trails to determine physical probabilities for particular events.This requires the processing of substantial quantities of data that may not available.

The alternate Bayesian approach uses the degree of a person's belief that an event will occur,rather than the actual probability that the event will occur. These probabilities are known as Bayesian probabilities and properties of person not the event.

The graphical representation of the dependencies embedded in the probabilistic models.The graphical representation used for belief propagation Techniques are two types ,they are Markov Random Fields and Bayesian Networks.[15]
Bayesian network use the directed graphs where the directions of the arrows permits distinguish genuine dependencies for spurious dependencies induced by hypothetical observations
In Markov Random Fields,the network topology was presumed to be given and problem was to characterize the probabilistic behavior of a system complying with dependencies prescribed by network.

Belief Propagation Techniques is used for performing inference on graphical models such as Bayesian Network and Markov Random Field[18] By using Belief Propagation Techniques wide variety of algorithms are developed in Artificial intelligence,Signal processing and Digital communication.[19]
The graphical representation of Markov Random Fields in Belief Propagation Techniques are used in first four applications

- **Efficient Belief Propagation for Early Vision.**[10]
- **Markov Network-based Unified Classifier for Face Identification**[5]
- **Efficient Loopy Belief Propagation using the Four Color Theorem**[6]
- **Image Completion Using Efficient Belief Propagation Via Priority Scheduling and Dynamic Pruning**[7]
- Belief propagation Algorithm is used in **Low Memory Cost Block-Based Belief Propagation For Stereo Correspondence**[2]

- Bayesian network method of graphical representation for **Task Parallel Implementation of Belief Propagation In Factor Graphs**[4]

- A graphical model such as the Markov random field (MRF) of Belief Propagation Technique is used in **Hardware-Efficient Belief Propagation**[8]

- A method of Particle Belief Propagation (PBP) is used in **PatchMatch Belief Propagation for Correspondence Field Estimation.**[**?**]

- Continuous time Bayesian network method is used in **Learning continuous time Bayesian network classifiers**[9]

The organization of report is as follows Introduction in **Chapter-1**,Mathematics related to Belief Propagation Techniques is in **Chapter-2**,Different applications related to Belief Propagation Techniques are in **Chapter-3**,conclusion in **Chapter-4.**

Chapter 2

Mathematics related to Belief Propagation Technique

2.1 Probabilistic Formulation:

Bayesian Methods provide a formalism for reasoning about partial beliefs under condition of uncertainty.. In the Bayesian formalism,beliefs measures obey the three basic axioms of probability theory.

$$0 \leq P(A) \leq 1 \tag{2.1}$$

$$P(SurePropositions) = 1 \tag{2.2}$$

$$P(AorB) = P(A) + P(B) if AandBaremutuallyexclusive \tag{2.3}$$

The basic expressions in the Bayesian formalism are statement about conditional probabilities be

The probability of any event A computed by conditioning it on any set of exhaustive and mutually exclusive events Bi,i=1,2,......n

$$P(A) = \sum\{P(A \setminus Bi) \times P(Bi)\} \tag{2.4}$$

This decomposition provides the basics for the hypothetical or assumption based reasoning in the bayesian formalism.It states that the beliefin any event A might be realized.

2.2 Probability Model

A probability model is an encoding of the probabilistic information in some domain of intrest,such that each well formed sentence,statement or proposition can be calculated in accordance with axioms of probability theory. These sentences are represented by interrelating random or stochastic variables with in model. The probability model (M) is defined by the joint probability of all its variables or universe

Each variable in a probabilistic model may take any one of a set of mutually exclusive or collectively exhaustive states. The random variable may be discrete ot continuous.The discrete random variable has finite of possible states,the probability distribution of Discrete random variable is probability mass function whereas continuous random variable takes their state from an infinite set of possible values.the probability distribution of continuous random variable is probability density function.

2.3 Probability Language

Probability language is well suited to reasoning under uncertainty.It provide a suitable frame work for processing the uncertainty relationship between variables of a probabilistic model. probability language provide consistency of inferred results and knowledge base through their axiomatic basis and provide a convenient mech anism for presenting uncertain results. The basic four primitives of probability language are

1. Likelihood
2. Conditioning
3. Relevance
4. Causation

- A likelihood of event is measure of how likely or probable it is that the event will occur ie.chance of occurrence.
- A conditioning :An event is conditional second event ,if its is changed by the knowledge of second event states.
- Relevance: events A and B are said to be relevant to each other in the context of event C ,if adding the knowledge of C to the knowledge of B,changes the likelihood of A. Example: Two events are relevance,when common sequences is observed.ie A and B are relevant if both are causes of C.
- Causation conveys a pattern of dependence between events.

2.4 Graphical Representation

A probabilistic model is dependency model in which the relationship between each of the variable is captured. A graph is denoted by G(V,E) is set of nodes or vertices V connected by the set of arcs or edge E. Graphs may be used to represent dependency model For example :Dependency Model M comprising the variables U represented by graph The set of arc E represents conditional dependencies between these variables. ie.Joint probability function of M is encoded in E

Nodes of the graph corresponds to variables in dependency model U$\rightarrow$ variables

Graphical representation for probabilistic model are two types one is undirected graph and directed graph. Undirected graph had no explicit direction,an arc of influence between connected nodes Examples of undirected graph is Markov Random Fields. In directed graph arcs are either unidirectional or bidirectional directed arc provide the mechanism for representing causation Example for Directed graph is Bayesian network

Markov Random Fields and Bayesian network are two types of graphical representation for probabilistic models

2.5 Bayesian Belief Network

The Bayesian Belief Network(BBN)determines the state probabilities of each node or variable from predetermined conditional and prior probabilities. The Direct Acyclic Graph(DAG) G(U,E)of the probability model M represents probability distribution P(U) ,where U represents set of all variables in M

$$X1, X2, ...Xn \tag{2.5}$$

Variables in probability distribution P(U) Direct Acyclic Graph(DAG) in which minimal set of variables are designed as parent of each variable Xi such that

$$P(xi/W); W \in X1, X2.....Xi - 1 \tag{2.6}$$

The above equation is BBN of that probability distribution For DAG to be Bayesian network of M,it is necessary and sufficient that each variable be conditional independent of all of its non descendents given in parents also satisfies this condition.
In practice usually understands the constrains in the domain of interest.However easily identify the variables that directly influence other variables ie.eaiser to understand the local interaction of variables than the domain as a whole.
The basic concept in the Bayesian treatment of uncertainties in causal network is conditional probability ie. Given the event B,the probability of event A is x

$$P(a/b) = x \tag{2.7}$$

Conditional probability

$$P(a/b)P(b) = P(a, b) \tag{2.8}$$

In real world all events are conditioned by some context C=c

$$P(a/b, c)P(b/c) = P(a, b/c) \tag{2.9}$$

The Baye's rule also conditioned on

$$P(b/a, c) = P(a/b, c)P(b/c)/P(a/c) \tag{2.10}$$

- P(b/a,c) is posterior probability of b
- P(a/b,c) liklihood probability
- P(a/c) Normalized factor
- P(b/c) Priori probability of b

Normalised factor is not directly availableis replaced by

$$\int_b P(a/b, c)p(b/c)db \tag{2.11}$$

$$\sum_b P(a/b, c)p(b/c) \tag{2.12}$$

For continuous and discrete distributions

2.6 Markov Field Belief Technique

The theory of Markov fields provides a safe method for constructing a complete and consistent quantitative model while preserving the dependency structure for an arbitrary graph (G).

The method consists of four steps

1. Identify the associates of G ,namely, the maximal subgraphs whose nodes are all adjacent to each other

2. For each associate Ci,assign a nonnegative compatibility function gi(ci) ,which measures the relative degree of compatibility associated with each value assignment cito the variable included in Ci

3. Form the product
   prodigi(ci) of the compatibility functions over all associates.

4. Normalize the product of all possible value combinations of the variables of the system

$$P(xi, ..., xn) = K \prod_{i} gi(ci) \tag{2.13}$$

where

$$K = 1 \div [\sum_{x1,...xn} \prod_{i} gi(ci)] \tag{2.14}$$

Chapter 3

Applications of Belief Propagation Techniques

Belief Propagation Technique is used for performing inference on graphical models such as Bayesian Network and Markov Random Fields.
Belief Propagation Technique calculates marginal distribution for each unobserved node, conditioning on any unobserved node.
Some of the applications where Belief Propagation Techniques are used explained below.

3.1 Efficient Belief Propagation for Early Vision

Markov random field models provide a robust and unified framework for early vision problems such as stereo and image restoration. Inference algorithms based on graph cuts and belief propagation have been found to yield accurate results, but despite recent advances are often too slow for practical use.
In this paper we present some algorithmic techniques that substantially improve the running time of the loopy belief propagation approach.

1. One of the techniques reduces the complexity of the inference algorithm to be linear rather than quadratic in the number of possible labels for each pixel, which is important for problems such as image restoration that have a large label set.

2. second technique speeds up and reduces the memory requirements of belief propagation on grid graphs.

3. A third technique is a multi-grid method that makes it possible to obtain good results with a small fixed number of message passing iterations, independent of the size of the input images.

Taken together these techniques speed up the standard algorithm by several orders of magnitude. In practice results obtained are as accurate as those of other global methods (e.g., using the Middlebury stereo benchmark) while being nearly as fast as purely local methods.
In this research paper concluded that three algorithmic techniques for speeding up the belief propagation approach for solving low level vision problems formulated in terms of Markov random fields.
The main focus of the paper is on the max-product formulation of belief propagation, and

the corresponding energy minimization problem in terms of costs that are proportional to negative log probabilities.
The first of the three techniques reduces the time necessary to compute a single message update from O (k 2) to O (k), where k is the number of possible labels for each pixel for the max-product formulation.

There are several opportunities for further development of these techniques.

- First, a general method for computing the min convolution quickly, analogous to the FFT for convolution, would broaden the applicability of fast message updates to arbitrary discontinuity cost functions based on difference between labels
- Second, the lower envelope method that we have presented for the min convolution could be extended to handle problems where the labels are embedded in some space but do not lie on a regularly spaced grid. More generally, it would be interesting to consider whether other sorts of structures on the set of labels enable fast methods.

3.2 Markov Network-based Unified Classifier for Face Identification

In this research paper ,a novel unifying framework using a Markov network to learn the relationship between multiple classifiers in face recognition is explored.
A several complementary classifiers and assign observation nodes to the features of a query image and hidden nodes to the features of gallery images. each hidden node is connected to its corresponding observation node and to the hidden nodes of other neighboring classifiers.
For each observation-hidden node pair, to collect a set of gallery candidates that are most similar to the observation instance, and the relationship between the hidden nodes is captured in terms of the similarity matrix between the collected gallery images. Posterior probabilities in the hidden nodes are computed by the belief-propagation algorithm.
The novelty of the proposed framework is the method that takes into account the classifier dependency using the results of each neighboring classifier. The two different evaluation protocols, known and unknown image variation tests, using three different databases, which shows that the proposed framework always leads to good accuracy in face recognition.

When a face image is used as as a query, can retrieve several desired face images from a large image database,then calculate many similarities of the query image and the gallery images in the database, and the retrieved gallery images are ranked by similar orders.
It is a one-to-many identification problem and has many applications such as searching similar face images in a database and face tagging in images and videos.
Recent successful face recognition methods have attempted to merge several classifiers using multiple feature sets of different characteristics, as in component-based methods, which extract features from separate spatial regions , and heterogeneous feature-based

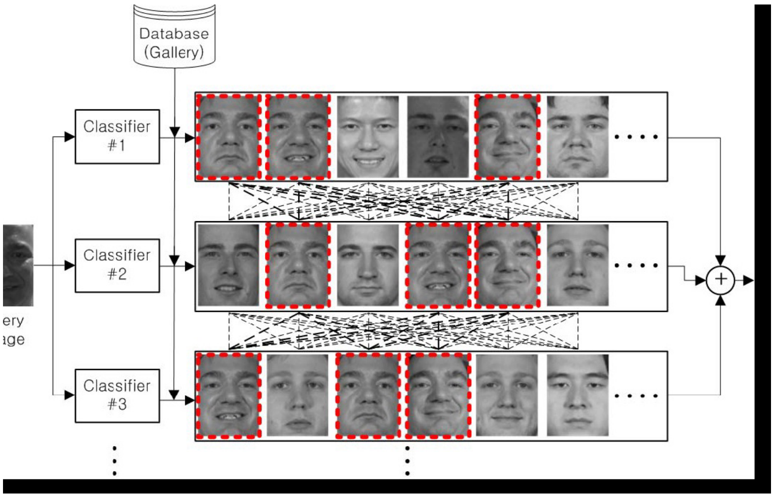

Figure 3.1: one-to-many identification

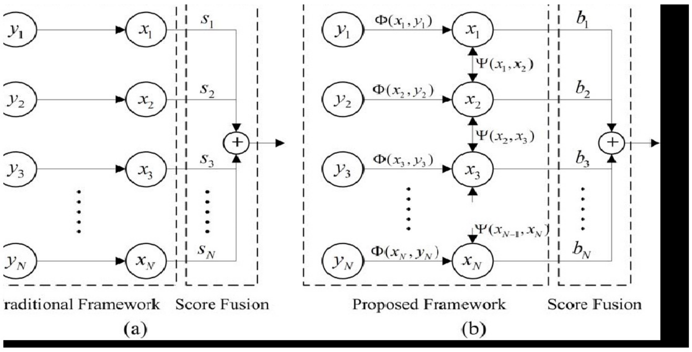

Figure 3.2:

methods which merge different domain features. These methods used the classifiers not only based on the different feature sets but also trained independently, and the similarity scores are merged with the predefined parameters. The parameter comes from the training database and it is not the best choice when the input image has different conditions. Note that these methods lead to good accuracy in face verification, but there is no specific framework for the one-to-many identification problem.

- In this paper, a novel recognition framework for the one-to-many identification issue is designed, and the simple concept is illustrated in Figure 3.1.
- First, assume that multiple classifiers have complementary characteristics, unify the multiple classifiers based not on the predefined weight values but on a Markov network, as summarized in Figure 3.2

For this purpose, assign one node of a Markov network to each classifier.The steps to find marginal probability by using Markov Fields for classifiers are as follows

1. Nodes are connected by lines, which represents the statistical dependencies

2. For an observation node, we extract a feature from a query image using the corresponding classifier.

3. At its paired hidden node, the first retrieve n similar gallery samples from the database, and their orders are made by the similarity scores for the query face image.

4. The multiple classifiers have their own lists of retrieved gallery images, which are not identical in general, thereby complementing the neighbor classifiers. Because the hidden nodes are connected by the network lines,

5. The relationship of the connected nodes is learned by the similarity scores between the neighbor classifiers, and the scores are calculated by concatenating the two gallery features of the neighbor classifiers.

6. The posterior probability at each hidden node is easily computed by the belief-propagation algorithm. Finally, marginal probability for a score value at each classifier is calculated

And also analyze the generalizability of the method using different multiple classifiers such as the Random Sampled Gabor (RSG) method, which consists of the simple and weak classifiers, and the Extended Curvature Gabor (ECG) method which consists of more complex and stronger classifiers.

Conclusion of this paper , A novel face recognition framework, particularly for the one-to-many identification task, based on multiple classifiers gallery connected by a Markov network.
The Markov network probabilistically models the relationships between a query and images and between neighboring gallery images.
From the viewpoint of an observation-hidden node pair, retrieve the most similar gallery images from the database using a query image face model.
The statistical dependency between the hidden nodes is calculated by the similarities between the retrieved gallery images.
Hence, the resulting inference mechanism can be viewed as a kind of clustering-based face recognition.

3.3 Efficient Loopy Belief Propagation using the Four Color Theorem

Recent work on early vision such as image segmentation, image denoising, stereo matching, and optical flow uses Markov Random Fields.
Although this formulation yields an NP-hard energy minimization problem, good heuristics have been developed based on graph cuts and belief propagation.
Nevertheless both approaches still require tens of seconds to solve stereo problems on recent PCs. Such running times are impractical for optical flow and many image segmentation and denoising problems and review on recent techniques for speeding them up.

Moreover in this research paper it shows that how to reduce the computational complexity of belief propagation by applying the Four Color Theorem (FCT) to limit the maximum number of labels in the underlying image segmentation to at most four. This provides substantial speed improvements for large inputs, and this for a variety of vision problems, while maintaining competitive result quality.

In this research paper two methods are developed based on graph cuts and belief propagation .

- In the case of Belief Propagation (BP), a key reason for its slow performance is that the algorithm complexity is proportional to both the number of pixels in the image, and the number of labels in the underlying image segmentation which is typically high. If limit the number of labels, its speed performance should improve greatly.
- By modifying the propagation algorithms like can using a low number of placeholder labels, that can reuse for non-adjacent segments. These placeholder labels can then be replaced by the full set of actual labels.
- Since image segments form a planar graph, they therefore require at most four placeholder labels by virtue of the Four Color Theorem (FCT) to still have different colors for all adjacent segments.
- A joint optimization process provides a fast segmentation through the placeholder labels and a fine grained labeling through the actual labels.
- The computational time is basically dependent on the number of placeholder rather than actual labels.

Statement of Four Color Theorem (FCT) is that

- for any 2D map there is a four-color covering such that contiguous regions sharing a common boundary (with more than a single point) do not have the same color
- The consequence of this theorem is that when an image, seen as a planar graph, is segmented into contiguous regions, there are only four colors to be assigned to each pixel/node for all segments to be surrounded only by segments of different colors .

Once such 4-color scheme is adopted, for each pixel/node there is only one of four decisions that can be taken.

The research work explored by using Four Color Theorem (FCT)are as follows

- This work exploits the FCT result to substantially improve the running time of BP, thus providing fast alternatives to local methods for early vision problems.
- Approach assigns one of 4 colors, i.e. one of 4 placeholder labels, to each pixel, in order to arrive at a stable segmentation of the image.At the same time it assigns a more fine-grained label, like the intensities, disparities, or displacements to each of the 4 possible colors.
- The resulting fine-grained labeling , the actual outcome of the algorithm changes continuously within the segments and abruptly across their boundaries.

In doing so, this approach provides a fast approximation to optimal MRF labeling. Henceforth Efficient Loopy Belief Propagation using the Four Color Theorem 3 systematically refer to placeholder labels as colors, and to actual, fine-grained labels as labels

Conclusion of the paper how the Four-Color Theorem based on the max-product belief propagation technique can be used in early computer vision for solving MRF problems where an energy is to be minimized. Methods used in this research yield results that are comparable with other methods, but improve either the speed for large images and/or large label sets (the case of image segmentation, stereo matching and optical ow), or both the performance and speed (the case of image denoising).

The Four Color Theorem principle is difficult to apply in cases where the label set is discrete and no natural order/relation between them can be inferred. This is the case for stereo matching and optical flow, where the disparity cost function takes discrete, unrelated values. This causes slower convergence, but is compensated by the low time complexity of the methods, independent of the number of labels. Thus, the proposed methods perform faster than the standard methods considered here, at least for large inputs.

3.4 Image Completion Using Efficient Belief Propagation Via Priority Scheduling and Dynamic Pruning

In this paper, a new exemplar-based framework is presented, which treats image completion, texture synthesis, and image inpainting in a unified manner. In order to be able to avoid the occurrence of visually inconsistent results, all of the above image-editing tasks in the form of a discrete global optimization problem. The objective function of this problem is always well-defined, and corresponds to the energy of a discrete Markov random field (MRF). For efficiently optimizing this MRF, a novel optimization scheme, called priority belief propagation (BP), is then proposed, which carries two very important extensions over the standard BP algorithm:

1. **Priority-based message scheduling**
2. **Dynamic label pruning**

These two extensions work in cooperation to deal with the intolerable computational cost of BP, which is caused by the huge number of labels associated with MRF. Moreover, both of extensions are generic, since they do not rely on the use of domain-specific prior knowledge.They can, therefore, be applied to any MRF, i.e., to a very wide class of problems in image processing and computer vision.
Thus managing to resolve what is currently considered as one major limitation of the BP algorithm: its inefficiency in handling MRFs with very large discrete state spaces. Experimental results on a wide variety of input images are presented, which demonstrate the effectiveness of our image-completion framework for tasks such as object removal, texture synthesis, text removal, and image inpainting.

THE problem of image completion can be loosely defined as follows: given an image which is incomplete, i.e., it has missing regions (e.g., see Fig. 1), try to fill its missing parts in such a way that a visually plausible outcome is obtained at the end. Although stating the image completion problem is very simple, the task of actually trying to successfully solve it, is far from being a trivial thing to achieve. Ideally, any algorithm that is designed to solve the image completion problem should have the following characteristics:

Figure 3.3: Object removal
Object removal is just one of the many cases where image completion needs to be applied. In the specific example shown, the user wants to remove a person from the input image on the left. He, therefore, simply marks a region around that person and that region must then be filled automatically so that a visually plausible outcome is obtained.

1. it should be able to successfully complete complex natural images
2. it should also be able to handle incomplete images with (possibly) large missing parts
3. all these should take place in a fully automatic manner, i.e., without intervention from the user.

Also, ideally, we would like any image completion algorithm to be able to handle the related problem of texture synthesis, as well. According to that problem, given a small texture as input, we are then asked to generate an arbitrarily large output texture, which maintains the visual characteristics of the input [e.g., see Fig. 2(a)]. It is exactly due to all of the above requirements that image completion is, in general, a very challenging problem. Nevertheless, it can be very useful in many areas, e.g., it can be important for computer graphics applications, image editing, film postproduction, image restoration, etc. It has, thus, attracted a considerable amount of research over the last years. There have been three main approaches so far, for dealing with the image completion problem [see Fig. 2(b)]:

1. Statistical-based methods
2. PDE-based methods
3. Exemplar-based methods

Exemplar-Based Methods Finally, the last class of methods consists of the so-called exemplar-based techniques, which actually have been the most successful techniques up to now. These methods try to fill the unknown region simply by copying content from the observed part of the image. All exemplar-based techniques for texture synthesis that have appeared until now, were either pixel-based or patch-based, meaning that the final texture was synthesized one pixel, or one patch at a time (by simply copying pixels or patches from the observed image, respectively).

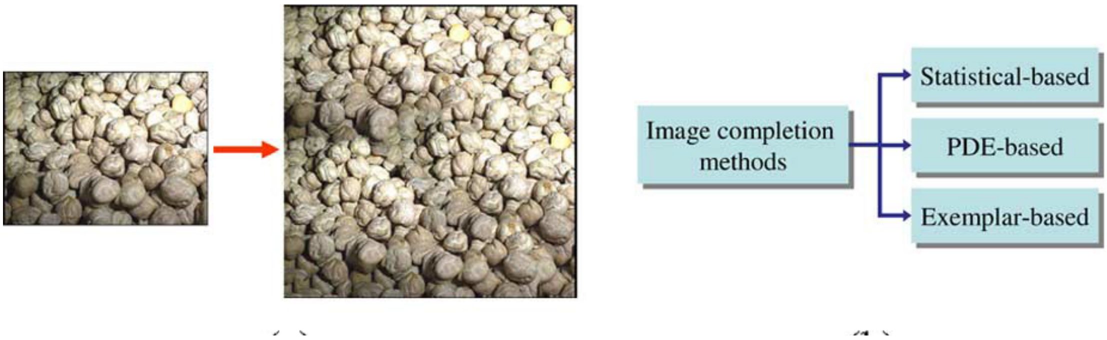

Figure 3.4: Texture synthesis problem and The three main approaches to image Completion

Conclusion of this research work A novel optimization scheme, priority-BP, has been proposed, that carries two very important extensions over standard BP: priority-based message scheduling and dynamic label pruning. This optimization scheme does not rely on any image-specific prior knowledge and can, thus, be to all kinds of images. Furthermore, it is generic (i.e., applicable to any MRF energy) and, thus, copes with one of the main limitations of BP: its inefficiency to handle problems with a huge number of labels.

Future work One of interesting avenue of future work would be to extend our framework so that it can be used for other types of completion problems, as well, e.g., it would be interesting to test framework on problems such as video completion or geometric completion. Also plan to allow the inclusion of more refinement terms into our energy function. Furthermore, this would make our method suitable for problems such as, e.g., constrained texture synthesis. Finally, besides image completion, we also plan to test our priority-BP algorithm, which is a generic MRF optimization scheme, to other labeling problems, for which the large cardinality of their state-space causes them to have a very high computational cost.

3.5 Low Memory Cost Block-Based Belief Propagation For Stereo Correspondence

The typical belief propagation has good accuracy for stereo correspondence but suffers from large run-time memory cost .
In this paper,It is proposed that a block-based belief propagation algorithm for stereo correspondence that partitions an image into regular blocks for optimization.
With independently partitioned blocks, the required memory size could be reduced significantly by 99

Stereo correspondence is used in computer vision to find the depth among the cameras and objects.
This depth inference problem could be further transformed to a disparity inference problem by assuming that the cameras and objects are under epipolar geometry. The inferred disparity information could be widely applied to tracking, surveillance system, and multiview video coding

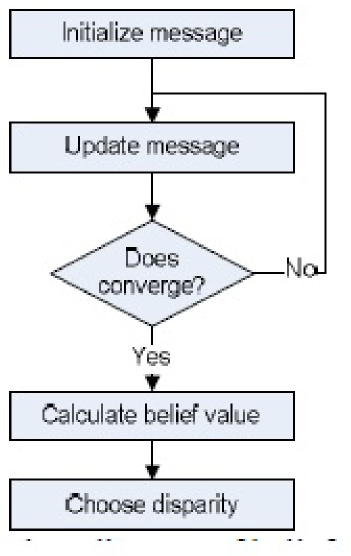

Figure 3.5: Flow diagram of Belief Propagation

The stereo matching algorithms can be roughly divided into two categories.

- **local approaches**
- **global approaches**
- Local approaches select disparities of image pixels using the information in a window. Therefore local approaches are faster than global approaches. However, it results in poor accuracy since the local approaches could not deal with textureless regions and occluded regions well due to the insufficient information in window.
- On the other hand, global approaches can handle the textureless and occluded regions well by formulating disparity inference as an energy minimization problem
- The energy function usually has a smoothness constraint which represents a certain physical relationship between neighboring pixel pair.
- This smoothness constraint often enforces penalty on the energy function, if the labels (disparities or segments) of neighboring pixels are inconsistent.
- Among the global methods, 2-D optimization algorithms such as graph cut and belief propagation (BP) have been applied quite successfully to optimize energy function.

The BP algorithms construct 2-D graph structures with nodes representing all the pixels in the disparity images to find the disparity map with energy closer to the global

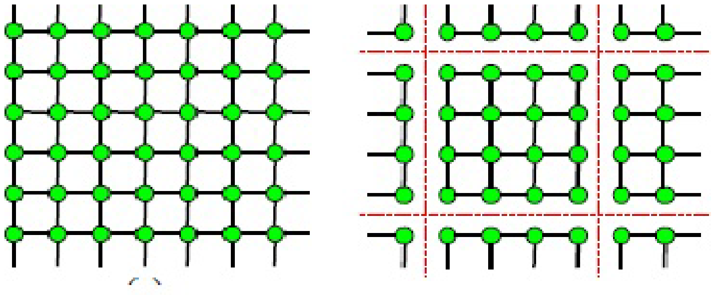

Figure 3.6: graph of typical and block based BP
2D Graph Model For Belief Propagation

minima. However, the vast number of nodes in the 2-D graph result in extremely high computation complexity, thereby rendering 2-D optimization is too difficult to be directly implemented for real-time application.
To address above problems, in this paper it is proposed that ,a block-based BP algorithm that directly partitions an image into separated independent blocks. Thus, can reduce the memory size significantly due to block based computation. In addition, the independent blocks also enable parallel computation by multiple computation units. Moreover earlier convergence for each block can also improve the long running time.

Conclusion of this paper,a new stereo matching algorithm partitions an image to block and optimizes with belief propagation technique.This method reduces memory storage size by 99 percentage with good performance.

Future work is possible is to enhance the interaction between neighboring blocks such that the independent block could extract useful information from neighboring finished processing blocks.

3.6 Task Parallel Implementation Of Belief Propagation In Factor Graphs

Factor graphs have been increasingly used as probabilistic graphical models. Belief propagation is a prominent algorithm for inference in factor graphs. Due to the high complexity of inference, parallel techniques for belief propagation are needed.
In this paper, the research work explored are task parallelism for belief propagation in an acyclic factor graph.The approach of this paper consists of building a task dependency graph based on the input factor graph and then using a dynamic task scheduler to exploit task parallelism system using a variety of acyclic factor graphs.

Graphical models have been essential tools for probabilistic reasoning. Factor graphs have emerged as a unified model of directed graphs (e.g. Bayesian networks) and undirected graphs (e.g. Markov networks).
A factor graph naturally represents a joint probability distribution that is written as a

product of factors, each involving a subset of random variables. Factor graphs have found applications in a variety of domains such as

- Image processing
- Bioinformatics
- Error-control decoding used in digital communications
- Inference is the problem of computing posterior probability distribution of certain variables given some value-observed variables as evidence.
- In factor graphs, inference proceeds with the well-known belief propagation algorithm . Belief propagation is a process of passing messages along the edges of a graph. Processing each message requires a set of operations with respect to the probability distribution of the random variables in a graph.
- Such distribution is represented by potential tables. The complexity of belief propagation increases dramatically as the number of states of variables and node degrees of a graph increase.

In many applications, such as digital communications, belief propagation must be performed in real time. Therefore, parallel techniques are needed to accelerate the inference.Many parallel techniques have been proposed for belief propagation in factor graphs. In this research techniques are developed for loopy belief propagation in cyclic factor graphs with the employment of embarrassingly parallel algorithms or with the need of graph partitioning.
Parallelizing belief propagation in acyclic factor graphs still remains a challenging problem due to the precedence constraints among the nodes in the graphs. In the meanwhile, task scheduling has been extensively studied and used in parallel computing task scheduling is shown to be an efficient tool for a class of linear algebra problems, known as regular applications, on general-purpose multi-core processors.
Since belief propagation in acyclic factor graphs is an irregular application, task scheduling is even a more suitable tool for parallelizing it.

- In this research work defining a task dependency graph for belief propagation
- then using a dynamic task scheduler to exploit task parallelism available in the task dependency graph.

conclusion of this research paper is

- The work on task parallelism for belief propagation in acyclic factor graphs is explored
- The method of approach used are constructing a task dependency graph for the input factor graph and then using a task scheduler to allocate tasks to the cores for parallel execution.

As part of future work is possible to plan to integrate data parallel technique to above mentioned method so that a large task can also be parallelized by the scheduler. In addition, the scheduler can group small tasks to reduce communication overhead and increase data locality.

3.7 Hardware-Efficient Belief Propagation

Loopy belief propagation (BP) is an effective solution for assigning labels to the nodes of a graphical model such as the Markov random field (MRF),

- But it requires high memory, bandwidth, and computational costs.

  The loopy BP has been widely applied to

- Stereo matching
- Image denoising
- Image inpainting

The success of BP is due to its regularity and simplicity. It uses a simple message update process to iteratively refine the beliefs of labels for each node. A message sent from one node to another is updated according to neighboring messages and local energy functions, using simple arithmetic operations.
However, BP algorithms generally require a great amount of memory for storing the messages, typically on the order of tens to hundreds times larger than the input data. Besides, since each message is processed hundreds of times, the saving/loading of messages consumes considerable bandwidth. Therefore, although BP may work on high-end platforms such as desktops, it cannot be applied to most consumer electronic devices that have limited memory, computational power, and energy. sequential procedure, it is difficult to utilize hardware parallelism to accelerate BP. In this paper, it is proposed that two techniques,

- The first one is tile-based BP and fast message construction, to address these issues. Tile-based BP splits the Markov random field (MRF) into many tiles and only stores the messages across the neighboring tiles. The memory and bandwidth required by this technique is only a fraction of the ordinary BP algorithms. But the quality of the results, as tested by the publicly available Middlebury MRF benchmarks, is comparable to other efficient algorithms.
- Second technique is that The fast message construction technique is based on the observation that many hypotheses used to construct the mesasages are repetitive. therefore, they only need to be computed once. This observation allows us to reduce the complexity of message construction from

Moreover, unlike previous sequential algorithms, the proposed algorithm can be easily parallelized. The proposed techniques can be realized in both hardware and software. A software reference implementation compatible to the Middlebury MRF library is available online, while two hardware examples [the first one is a very large scale integration (VLSI) circuit and the second one a graphic processing unit (GPU) program are analyzed in this paper.

Conclusion The techniques used to develop a tile-based message passing and fast message construction algorithm greatly reduced the memory, bandwidth, and computational costs of BP and enabled the parallel processing. With these two techniques, BP becomes more suitable for low-cost and power- limited consumer electronics. These techniques can be applied to other parallel platforms

3.8 PMBP: PatchMatch Belief Propagation for Correspondence Field Estimation

Patch Match is a simple, yet very powerful and successful method for optimizing Continuous labelling problems. The algorithm has two main ingredients:

- the update of the solution space by sampling
- the use of the spatial neighborhood to propagate samples.

These ingredients are related to steps in a specific form of belief Propagation in the continuous space, called Particle Belief Propagation (PBP). However, BP has thus far been too slow to allow complex state spaces.
The two approaches used in this research yields a new algorithm called PMBP Patch Match Belief Propagation for Correspondence Field estimation, which is more accurate than Patch Match and orders of magnitude faster than PBP. The methods used in research is novel realistic pair wise terms that provide Smoothness used for the recent Patch Match Stereo work.

Conclusion of this research is the link between the popular PatchMatch method and the very well-known Belief propagation algorithm. By doing so able to extend the Patch-Match algorithm by introducing additional pairwise terms

Future work There are many exciting avenues for future work, both in terms of applications, such as optical flow, as well as algorithms, such as different forms of message passing e.g. Treereweighted message passing

3.9 Learning continuous time Bayesian network classifiers

Streaming data are relevant to use in finance,computer science,engineering while they are becoming increasingly important to medicine and biology.

The research work on this paper are as follows:

- Continuous time Bayesian network classifiers are designed for analyzing multivariate streaming data when time duration of event matters.
- Structural and parametric learning for the class of continuous time Bayesian network classifiers are considered in the case where complete data is available.
- Conditional log-likelihood scoring is developed for structural learning on continuous time Bayesian network classifiers.
- Performance of continuous time Bayesian network classifiers learned when combining conditional log-likelihood scoring and Bayesian parameter estimation are compared with that achieved by continuous time Bayesian network classifiers when learning is based on marginal log-likelihood scoring and to that achieved by dynamic Bayesian network classifiers.

- Classifiers are compared in terms of accuracy and computation time. Comparison is based on numerical experiments where synthetic and real data are used.
- Results show that conditional log-likelihood scoring combined with Bayesian parameter estimation outperforms marginal log-likelihood scoring. Conditional log-likelihood scoring becomes even more effective when the amount of available data is limited.
- Continuous time Bayesian network classi?ers outperform in terms of computation time and accuracy dynamic Bayesian network on synthetic and real data sets.

Conclusions Conditional log-likelihood scoring function has been developed to learn continuous time Bayesian network classifiers from multivariate streaming data.
A new learning algorithm for continuous time Bayesian network classifiers is designed by combining conditional log-likelihood scoring for structural learning with marginal log-likelihood parameter estimation.
Conditional log-likelihood scoring outperforms the marginal log-likelihood scoring in terms of the accuracy achieved by continuous time Bayesian network classifiers

Future work : Continuous Time Bayesian Network Classifiers efficiently address the problem of classifying multivariate trajectories in the case where the class is static and the multivariate trajectories are completely observable.

A possible future step is to study if it is possible to preserve the e?ciency of the classification algorithm, relaxing the necessity to work with completely observable trajectories.

The quality of the classification performances also suggests to extend the Continuous Time Bayesian Network Classifiers to the clustering problem.

Chapter 4

Conclusion

A complete probabilistic model for the variables and their relationship can be build by using Belief Propagation Techniques.

In Artificial intelligence,Signal processing and Digital communication stochastic or probabilistic models are often formulated by using Belief Propagation Techniques.

The graphical representation or models for multidimentional probability distributions such as Markov Random Fields and Bayesian Networks are used.

The basic concepts of Markov Random Fields and Bayesian Networks are studied

In recent research papers where Belief Propagation Techniques approach of solving various problems in domain of image processing in terms of using Markov Random Fields and Bayesian Networks are studied.

Bibliography

[1] Judea Pearl *Probabilistic Reasoning in Intelligent Systems* . KAUFMANN MORGANN publishers, second revised and pdated edition,San Francisco, California, 1998.

[2] yu-chang andNelson chang `LOW MEMORY COST BLOCK-BASED BELIEF PROPAGATION FOR STEREO CORRESPONDENCE`, 1-4244-1017-7187,IEEE Transisition

[3] F.BESSE and C.ROTHER,A.FITZGIBBON and J.KAUTZ `PMBP:PatchMatch Belief Propagation for Correspondence Field Estimation` University college of London U.K.,Microsoft Research Cambridge ,2012

[4] Nam Ma,Yinglong xia,Viktorprasanna `Task Parallel Implementation of Belief Propagation In Factor Graphs` This Research was suppored by U.S.National Foundation under grant no.CNS-101880

[5] Wonjun Hwang Kyungshik Roh Junmo Kim `Markov Network-based Unified Classifier for Face Identification` This ICCV2013 paper is open access available in IEEE Xplore

[6] Radu Timofte,Lucvan Gool `Efficient Loopy Belief Propagation using the Four Color Theorem` ISICS,ESAT-PSI/iMinds V,Belgium,Computer Vision Lab,Zurich ,Switzerland

[7] Nikos Komadakis and Georgios Tziritas `Image Completion Using Efficient Belief Propagation Via Priority Scheduling and Dynamic Pruning` IEEE Transactions in Image Processing ,volume;16,No.11,November 2007

[8] Chia-Kai Liang,Chao-Chung cheng,Yen-Cheieh Lai `Hardware-Efficient Belief Propagation` IEEE Transactions on circuits and systems for Video Technology Vol;21,No.5,May 2011

[9] Daniele Codecasa,Fabiostella `Learning continuous time Bayesian network classifiers` International Journal of Approximate Reasoning,55(2014) 1728-1746

[10] PEDRO F. FELZENSZWALB.DANIEL P.HUTTENLOCHER `Efficient Belief Propagation for Early Vision` International Journal of Computer vision (Springer Science)70(1),41-54,2006

[11] Dan Yuan `Understanding Belief Propagation and its Applicatons` Department of computer Engineering,University of California

[12] Erik.b.Sudderth and William T.Freeman `Signal and Image processing with Belief Propagation` Accepted to appear in IEEE Signal Processing Magazine

[13] N.Kmodakis and G.Tziritas `Image Completion using Efficient BP via priority Scheduling and dynamic Pruning` IEEE trans.in Image Process,vo.16 NO.11

[14] A,Brunton.C.Shu and G.Roth `Belief Propagation on GPU for Streo Vision` IEEE Proceedings CRV 2006.

[15] Hans Andrea Loeliger `An Introduction to Factor Graphs` IEEE Signal Processing Magazine,2004

[16] Jianguo Ding `Probabilistic Inferences in Bayesian Networks` International Journal of Computer science,arXiv:1011.0935v2,5 November 2010

[17] Frank R.Kschischang,Brendan J.Frey,Hans Andrea Loeliger `Factir Graph and the Sum -Product Algorithm` IEEE Transcations on information Theory,vol47.No.2,Feb,2001

[18] Judea Pearl `Reverend Bayes on inference Engines:A distributed Hierarchical Approach` AAAI-B2 Proceedings 1982

[19] Heckerman.D. `A Tutorial on Learning With Bayesian Networks` Technical Report MST-TR-95-96,Microsoft Corporation ,Redmond,WA